Chapter 5
A Cross-Scale Approach for Abundance Estimation of Invasive Alien Plants in a Large Protected Area

Cang Hui, Llewellyn C. Foxcroft, David M. Richardson, and Sandra MacFadyen

Abstract Efficient management of invasive alien plants requires robust and cost-efficient methods for measuring the abundance and spatial structure of invasive alien plants with sufficient accuracy. Here, we present such a monitoring method using ad hoc presence-absence records that are routinely collected for various management and research needs in Kruger National Park, South Africa. The total and local abundance of all invasive alien plants were estimated using the area-of-occupancy model that depicts a power-law scaling pattern of species occupancy across scales and a detection-rate-based Poisson model that allows us to estimate abundance from the occupancy, respectively. Results from these two models were consistent in predicting a total of about one million invasive alien plant records for the park. The accuracy of log-transformed abundance estimate improved significantly with the increase of sampling effort. However, estimating abundance was shown to be much more difficult than detecting the spatial structure of the invasive alien plants. Since management of invasive species in protected areas is often hampered by limited resources for detailed surveys and monitoring, relatively simple and inexpensive monitoring strategies are important. Such data should also be appropriate for multiple purposes. We therefore recommend the use of the scaling pattern of species distribution as a method for rapid and robust

C. Hui (✉) • D.M. Richardson
Centre for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Stellenbosch 7602, South Africa
e-mail: chui@sun.ac.za; rich@sun.ac.za

L.C. Foxcroft • S. MacFadyen
Conservation Services, South African National Parks, Private Bag X402, Skukuza 1350, South Africa

Centre for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Stellenbosch 7602, South Africa
e-mail: Llewellyn.foxcroft@sanparks.org; sandramf@live.co.za

L.C. Foxcroft et al. (eds.), *Plant Invasions in Protected Areas: Patterns, Problems and Challenges*, Invading Nature - Springer Series in Invasion Ecology 7,
DOI 10.1007/978-94-007-7750-7_5,

monitoring of invasive alien plants in protected areas. Not only do these approaches provide valuable tools for managers and biologists in protected areas, but this kind of data, which can be collected as part of routine activities for a protected area, provides excellent opportunities for researchers to explore the status of aliens as well as their assemblage patterns and functions.

Keywords Invasive plants • Kruger National Park • *Opuntia stricta* • *Lantana camara* • Optimum monitoring strategy

5.1 Introduction

Population size (abundance) is a measure of the extent and impact of biological invasions (Parker et al. 1999). Therefore, to assess the impact of invasive alien species and propose efficient and timely control strategies, we need to have a cost-efficient and robust approach for estimating abundance. Without this it is difficult to assess the status and change in invasions, and to propose refinements to existing management strategies. To maximise the application of this approach, we cannot only utilise datasets collected and recorded for specific purposes (e.g. relative abundances recorded along transects), but those in widely-accessed format collected in the field by conservation and alien plant control managers. To quickly assess the invasion extent, ground managers often accumulate localities of a focal invader found in a haphazard fashion or reported by others. To this end, designing methods that can utilise these geo-referenced presence-only or presence-absence records and robustly estimate the abundance at a satisfactory level could provide substantial benefits and opportunities for control and risk assessment, and thus deserves closer attention.

Many types of models can be used to estimate abundance from presence-absence records (e.g. Wright 1991; MacKenzie et al. 2002; He and Gaston 2003; Royle et al. 2005; Hui et al. 2009, 2011a). These models can be clustered into three groups. First, models based on the abundance-occupancy relationship require a robust representation of the focal species' occupancy based samplings of spatially equal effort (e.g. He and Gaston 2003). These types of models often assume that the environment is homogenous (Hui et al. 2009) and thus can only be used at local scales. Second, models incorporating imperfect detection rate often require repeated samples (MacKenzie et al. 2002; Hui et al. 2011a); this often not feasible when working over large areas. Moreover, using this type of model, and also interpreting the model predictions, usually requires knowledge of fairly advanced mathematical methods. This has hindered their wide application in the field. Moreover, both types of models require systematic sampling and are only appropriate for studies of particular species at local sites. They fall short when a large-scale quick assessment is needed.

We here recommend the use of a third kind of model, namely the area-of-occupancy (AOO), which does not involve abstruse mathematics. They are also suitable for rapid, large-scale assessments using records. A power-law pattern of

AOOs across scales has been widely observed for species, across taxa (Kunin 1998). The power law AOO could reflect the fractal structures of species distributions driven by the fractal structures of environmental variables and multiple ecological processes. Although the exact reason behind a power-law AOO is still unknown and widely debated (e.g. Halley et al. 2004 and references therein), this specific form of AOO model has been suggested as an elegant way for extrapolating species occupancy across scales (Wilson et al. 2004) and estimating abundance with reasonable accuracy over large, heterogeneous environments (Hartley and Kunin 2003; Hui et al. 2009). Due to the lack of suitable large-scale monitoring approaches, incorporating the AOO model into monitoring programmes for alien and invasive plants is an attractive option. Here, we aim to (i) explore the presence records collected by rangers during their routine patrols in protected areas and (ii) demonstrate how the AOO method can help us to estimate the abundances and distributions of top plant invaders in the protected areas.

5.2 Kruger National Park as a Model System

We demonstrate the AOO approach for the abundance estimation of invasive alien plants (IAPs) in a large protected area, which is actively managed for wildlife protection and biodiversity conservation – South Africa's Kruger National Park (KNP). Kruger National Park (c. 20,000 km^2, the size of Israel) is situated in the north-eastern part of South Africa, in a semi-arid savanna system (Fig. 5.1a). A full account of the KNP's biophysical landscape and its rich 100-year management and research history has been given in Du Toit, Rogers and Biggs (2003). Unfortunately KNP also contains a large number of alien species, with 348 invasive alien plant (IAP) species identified to date (Foxcroft et al. 2003; Spear et al. 2011).

In 2004 an electronic handheld data capturing system – CyberTracker – was initiated. The CyberTracker system is a customised, icon driven programme run on a personal digital assistant (PDA) device, and is used during routine patrols by rangers, to capture distribution records of a number of features such as animal sightings, ephemeral water distribution, carcasses, alien species and many others (full details on the CyberTracker programme are given in Kruger and MacFadyen 2011). Examples of other uses of CyberTracker data are rich in literature (Dietemann et al. 2006; Foxcroft et al. 2009; Hui et al. 2011a). Approaches with which the abundance of IAPs may be estimated using the CyberTracker records are thus not only of management value for KNP, but also provide an opportunity to assess, refine and provide guidelines for designing similar monitoring programmes for other protected areas.

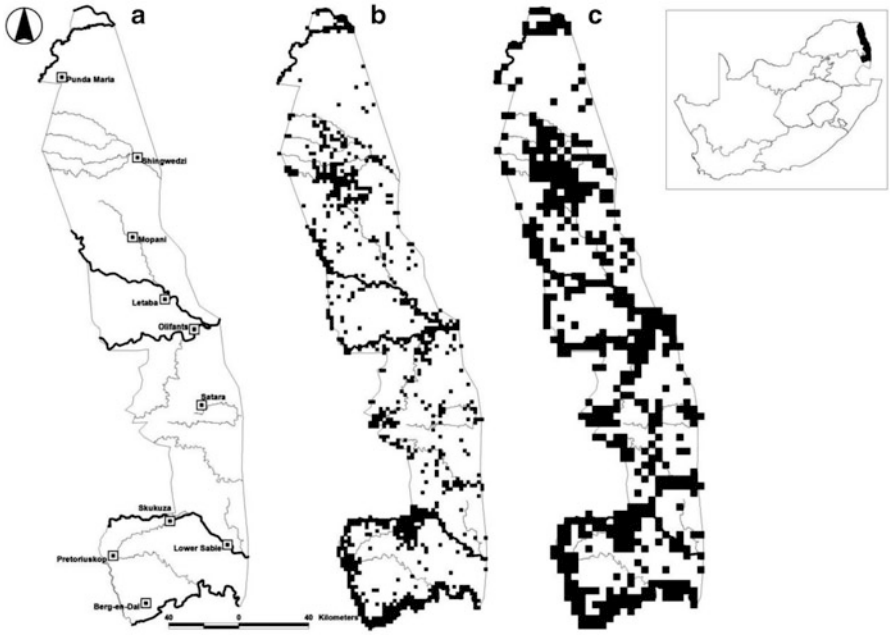

Fig. 5.1 (**a**) The spatial geography of Kruger National Park and presence-absence maps of invasive alien plants at a resolution of 2 × 2 km (**b**) and 4 × 4 km (**c**) of the Kruger National Park

5.3 Power Law Area-Of-Occupancy

To estimate the total number of all IAPs in the KNP, we first divided the landscape into lattices with the grain size of each cell a km^2 (Fig. 5.1b, c). Let $u_a(i)$ denote the number of records in the cell i, in which a number of $u_a^+(i)$ records are presence and the rest absence. The occupancy at the scale of a can thus be defined as (Eq. 5.1),

$$P_a = \sum_i a \cdot \varphi_a(i) \tag{5.1}$$

where $\varphi_a(i) = 1$ if $u_a^+(i) \geq 1$; otherwise, $\varphi_a(i) = 0$. The power-law AOO describes how the occupancy P_a changes across scales, specifically as a function of grain a (Kunin 1998; Hartley and Kunin 2003; Gaston and Fuller 2009; Hui et al. 2009). Therefore, there are a number of $N_a = P_a/a$ cells occupied by the IAPs at the scale of a; that is, the number of presences at the scale a is (Eq. 5.2),

$$N_a = \sum_i \varphi_a(i) \tag{5.2}$$

When the grain is small enough to only hold at most one IAP individual (i.e. the grain equals the individual size, $a = \delta$), the number of occupied cells N_δ is thus equal to the abundance of the IAPs (Hartley and Kunin 2003). The individual size is not the canopy size of a plant but equals the reciprocal of local density; that is, the minimum land size that can support the growth of a plant. For plant species that are unfeasible for identifying individuals, we can use the reciprocal of the density of plant patches to define δ; of course, N_δ will not represent the abundance but rather the number of plant patches of a focal species. A power-law form of the AOO has been confirmed for the distribution of plants (Kunin 1998; Hartley and Kunin 2003; Hui et al. 2011b) and butterflies (Wilson et al. 2004),

$$P_a = c \cdot a^d \tag{5.3}$$

where c and d are constants (Eq. 5.3). Parameter c represents the occupancy when the grain equals 1 km^2, whilst d denotes the exponent of the power law and is proportional to how fast the occupancy changes with the spatial scales when the grain is around 1 km^2. We chose this power-law AOO for the abundance estimation because of its simplicity and empirical support of plant distributions. The power-law AOO also indicates a self-similar and fractal nature of species distribution (Hui and McGeoch 2008).

The power-law AOO is a powerful predictive tool, as once the occupancies at two (or more) different scales have been determined, the parameters c and d can be estimated and therefore the total abundance N_δ can be calculated. For a robust calculation we divided the landscape into lattices with 64 different grains, with the width of the cells ranging from 125 m to 8 km. The power law AOO was fitted using linear regression on log-log transformed axes for the occupancies at these 64 scales. The number of occupied cells was also calculated for seven classes to understand the species-level assembly of IAPs: *Opuntia stricta* (sour prickly pear), *Lantana camara* (lantana), *Opuntia* spp. (all other *Opuntia* records combined, except those specified as *O. stricta*), *Chromolaena odorata* (Chromolaena/Siam weed), *Pistia stratiotes* (water lettuce), *Parthenium hysterophorus* (parthenium) and others (the combination of species with less than 200 records). Lennon et al.'s (2007) χ^2 test of the difference in maximum log-likelihoods was used to verify this power-law AOO.

To measure the performance of using the power-law AOO for total abundance estimation under different sampling efforts, we calculated the occupancies P_a and P_{4a} (by combining the four adjacent a-size cells to form a $4a$-size cell; specifically $a = 4 \times 4$ km and thus $4a = 8 \times 8$ km), from which the parameters c and d can be calculated and the total abundance can be estimated. The status (presence or absence) of boundary cells was determined solely on the records within the park. Since the occupancy of a $4a$-size cell could be derived from the combined information of less than four a-size cells (e.g. cells along the boundaries), the abundance

estimates at coarse scales were likely to be overestimated. We propose a general relationship between the abundance estimation and sampling effort (Eq. 5.4),

$$\ln N(s) = \ln N_{\delta} \cdot \left(1 - \exp\left(-a \cdot s^{\beta}\right)\right) \tag{5.4}$$

where $N(s)$ is the abundance estimation under the sampling effort s; α is a measure of the converging speed from the abundance estimation $N(s)$ to its limit N_{δ}; β is a scaling parameter for sampling effort. Let A = 1-Abs[pred-obs]/pred denote the accuracy of the prediction (Hui et al. 2006), the above *abundance-effort relationship* would allow us to estimate the minimum sampling effort (s^{*}) for estimating the abundance at a satisfactory level (say, $A = 0.95$).

5.4 Poisson Occupancy-Abundance Model

For the calculation of the local abundance, i.e. the abundance of IAP in specific cell, we first estimated the proportion of presences in each cell, $D_a(i) = u_a^{+}(i)/u_a(i)$. Second, if we assume that CyberTracker records occur randomly and independently of one another within the cell, we can estimate the local abundance of IAPs according to the Poisson occupancy-abundance model (Eq. 5.5; Wright 1991),

$$n_a(i) = -U \times \ln(1 - D_a(i)) \tag{5.5}$$

where $1/U$ is the minimum size of the area that one record can represent; that is, the size of the detection area when a ranger stands still. In practice, U is estimated by the size of grain divided by the maximum number of records in cells. The local abundance was calculated at a grain of 4 × 4 km for demonstration. More sophisticated approaches that deal with pseudo-absence dilemma and zero-inflation problem are also available (e.g. Bayesian estimation model; Hui et al. 2011a) but we prefer this simple Poisson model here that also provides a reasonable estimate of local abundance. The AOO method was not used for calculating the local abundance because additivity does not apply for the power law (e.g. $x^{1.5}+y^{1.5} < (x+y)^{1.5}$; Cohen et al. 2005) and thus the summation of local abundances often underestimate the total abundance.

To measure the performance of the above approach under different sampling efforts for estimating the spatial structure of IAPs, the normal approach would be to calculate the discrepancy between the frequency distributions of the predicted local abundance and observed local abundance, using chi-square statistics. However, these statistics encounter problems when the frequency for some categories is lower than five (Quinn and Keough 2002), which is often the case during the simulation. Instead, we use the sum of squared errors for the distance (or deviation) between the original spatial structure and the spatial structure from sampling (Eq. 5.6),

$$d = \sum_i \left(\log_2 D(i) - \log_2 D_a^{'}(i) \right)^2 \tag{5.6}$$

where $D_a^{'}(i)$ is the estimate of the proportion of presences in the cell i under a specific sampling effort. The sum and mean of squared errors have often been used in comparing the similarity of two images in the field of image processing (e.g. Wang et al. 2004). The reason for the log-transformation here is because it can largely normalise the observed frequency distribution of the proportion of presences.

The significant distance can be determined by a randomization test (also called a permutation test; Sokal and Rohlf 1995). For this test we reshuffled the observed proportions of presences for the cells with at least one presence record 5,000 times. After each run, we calculated the above distance between the reshuffled spatial structure and the original observed spatial structure. We then built a probability distribution of these 5,000 distances in order to proceed with a test of significance. We found that the observed probability distribution of these distance was not different from a normal distribution (Shapiro-Wilk test; SW-W = 0.9995, p = 0.217), from which we identified a one-tail critical value with p = 0.05, $d_{0.05}$ = 5,027.5. With certain amount of sampling effort, we can predict the spatial structure of local abundances of IAPs. The deviation of this sampling-effort-dependent structure from the real structure can be captured by d; if we have $d < d_{0.05}$, the predicted spatial structure is not significantly different from the observed structure ($p < 0.05$); otherwise ($d > d_{0.05}$), the spatial structure prediction is significantly different from that which was observed. The minimum sampling effort was defined as the sampling effort so that the spatial structure from simulations is not significantly different from the observed spatial structure.

5.5 Invasion Status of Kruger National Park

We used data from 2004 to 2007, comprising 2.4 million records with 27,777 presences and the rest absences (absences were taken from records of other features that were collected; see Hui et al. 2011a for issues with pseudo-absences). The extremely low occurrence (1.15 %) of IAPs does not necessarily indicate insufficient sampling intensity but could suggest that most species are currently: (i) at an early phase of invasion, with high potential for rapid expansion, (ii) are being maintained at the current state of low abundance through on-going management, or (iii), for some species at least, especially those introduced from regions with a completely different climate, have already occupied the full extent of their potential range in the area.

There were 167 IAP species represented by the presence records, with most records for *O. stricta* (72.1 %) and *L. camara* (8.4 %), and 119 species with fewer

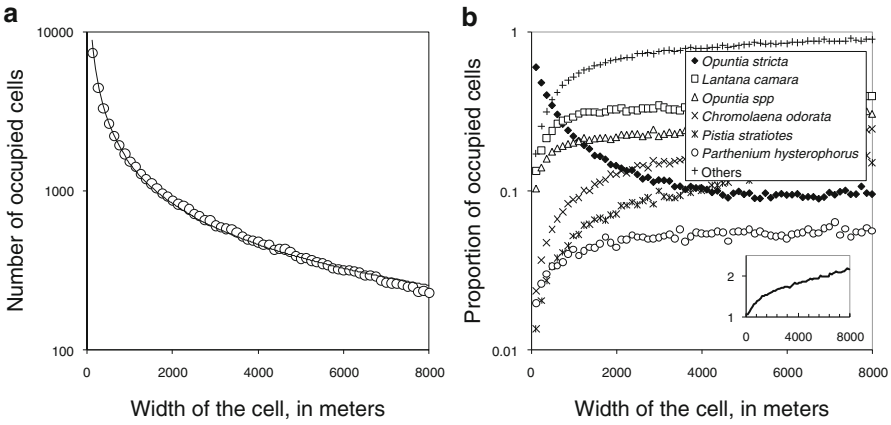

Fig. 5.2 (**a**) The scaling pattern of the number of occupied cells of all the invasive alien plants in Kruger National Park. The *solid line* indicates the regression from all 64 grains; the *solid circles* indicate the number of occupied cells at 64 spatial scales (from 125 × 125 m to 8 × 8 km). The *fitted lines* have been extrapolated over two orders of magnitude to the scale of 1 × 1 m. (**b**) The proportion of cells occupied by each species in the cells with records of IAPs. The inset indicates the sum of the proportion of all species approaches one when scaling down

than ten records. The maximum number of records per cell for the grain of 4×4 km^2 is $U = 72{,}203$, indicating a detection area of 221.6 m^2 with a diameter of 16.8 m. The mean canopy size of all IAPs weighted by the number of records for each species is 0.748 (0.743–0.754) m^2 (a mean radius of 0.49 m). For simplicity, we added a buffer zone beyond the canopy (an extension of 0.076 m) to make the individual size to be exactly 1 m^2 (i.e. the total abundance estimates equal the number of occupied unit-size cells).

The power-law form of the area-of-occupancy (AOO) provided an accurate description of the scaling pattern of IAP's distribution (Fig. 5.2a). The maximal log-likelihood (= −6,160.38) of Lennon et al.'s (2007) scale-dependent model showed no significant difference ($\chi^2 = 3.72$; $p > 0.05$) from the power law AOO ($\text{log-likelihood}_{\max} = -6{,}162.24$). When considering all 64 scales, we had $N_a = e^{13.2} a^{-0.425}$ ($R^2 = 0.995$; solid line in Fig. 5.2a), indicating a total number of 552,820 IAPs with a box-counting fractal dimension $D = 2 \times 0.425 = 0.85$. We further calculated the AOO for 16 smaller scales (from 125 × 125 m to 2 × 2 km; $N_a = e^{12.7} a^{-0.385}$, $R^2 = 0.999$, $D = 0.77$; a total number of 313,461 IAPs) and for 49 larger scales (from 2 × 2 km to 8 × 8 km; $N_a = e^{14.2} a^{-0.485}$, $R^2 = 0.97$, $D = 0.97$; a total number of 1,450,685 IAPs).

Species-level partitioning of the occupied cells showed a decrease of overlapping among species when scaling down, with the sum of occupied cells for each species approaching the number of occupied cells when combining all IAP species together (Fig. 5.2b), suggesting a reliable estimation of the total number of IAPs. Although the log-transformed abundances are compatible, disparities do appear among these estimates – a typical problem of any scaling

method (Lennon et al. 2007). Therefore, the abundance estimate must be verified by the sum of local abundance. When the width of the grain cell is greater than 4 km, the proportion of occupied cells for each species is fairly stable (i.e. scale-insensitive; Fig. 5.2b). However, it starts to group into three different levels at finer scales: *C. odorata*, *P. stratiotes* and *P. hysterophorus* converges to only occupying 1 % of the total occupied cells of IAPs; *L. camara*, *Opuntia* spp. and other species combined converges to occupy 10 % of the occupied cells; *O. stricta* converges to occupying a majority of all occupied cells (Fig. 5.2b). Evidently, most IAP species have a fairly wide range but are locally rare, whilst *O. stricta* has a moderate range but are locally extremely abundant.

We calculated the local abundance at the scale (grain) of 4 × 4 km. Because the Poisson model can only be applied for cells with presence records (Fig. 5.1), the estimation of local abundance were literally for those areas only. Abundance of IAPs in absence cells was thus considered extremely low and was neglected in the further analyses. The total abundance of IAPs calculated by summing up the local abundances of all cells is 1,033,969, which is half way between the AOO estimates from all scales (552,820) and from large scales (1,450,685).

5.6 Assemblage Patterns of Aliens

Preston (1948) first identified the log-normal form of the "species" and "individual" curves, i.e. the frequency distribution of species and individuals falling in any given octave class. Based on his finding, we expect a log-normal form of the frequency distribution of local abundance in the cells (i.e. the number of IAPs in each cell). Even though the $\log_2$-transformed frequency distributions of the proportion of presences and local abundance have similar shape (Fig. 5.3), neither followed a strict normal distribution (SW-W = 0.963, $p < 0.01$; SW-W = 0.957, $p < 0.01$; Fig. 5.3).

If we consider each cell to be identical in hosting individuals (i.e. a homogenous landscape), a neutral model prediction that considers the birth and death events within each cell, should be expected (Volkov et al. 2005). This is equal to switching the concept of a species' abundance in a community by the number of IAPs in a cell; that is, the number of individuals of a neutral species behaves similarly to the number of individuals of a homogenous landscape site. Therefore, we can test whether the frequency of cells with n number of IAPs (F_n) follows Volkov et al.'s (2005) neutral model, $F_n = \theta x^n/(n + c)$, where x and c are regression parameters, and θ a normalization constant to ensure the number of cells sensible). Parameter x indirectly represents the increase rate of IAPs, and c controls the strength of the density dependence. Using maximum log-likelihood method for parameterization, we also found a significant distinction between the neutral model prediction from the local abundance estimates (log-likelihood = −1599.1, $x = 0.99992$, $c = 33$, $\theta = 0.1865$, $\chi^2 = 75.8$, $df = 10$, $p < 0.01$).

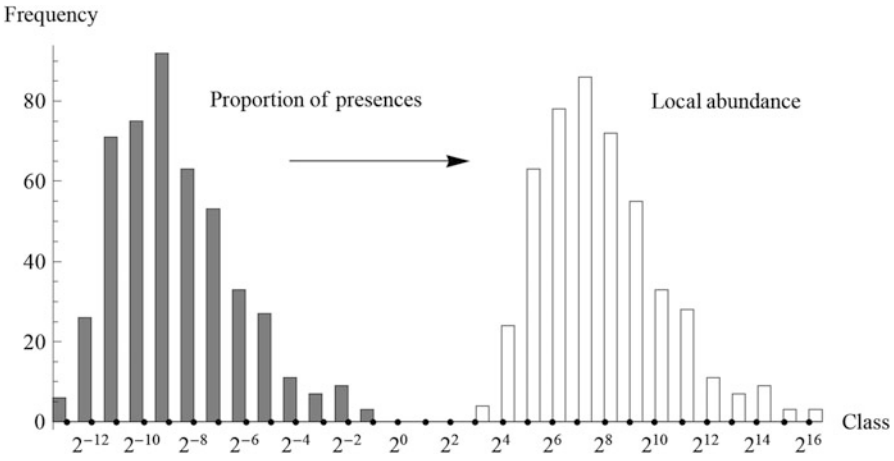

Fig. 5.3 Frequency distributions of the proportion of presences (*grey bars*) and the local abundance (i.e. the number of IAP individuals per cell; *white bars*) of the Kruger National Park at the 4 × 4 km scale

Evidently, besides the overall log-normal shape of the local-abundance frequency distribution (Fig. 5.3), it differs from both log-normal shape and neutral model predictions. This indirectly suggests (i) that the KNP landscape is heterogeneous (thereby supporting other work, for example Pickett et al. 2003) and (ii) that IAPs in KNP have not reached their demographic equilibrium, indicating that the invasions are at an early stage or that, in some areas at least, IAPs are being maintained at their current state through management efforts. Of course, other interpretation of this result may also exist. For instance, the maximum likelihood estimate of parameter c is much greater than those estimated for tropical forests (Volkov et al. 2005), suggesting there could be a strong density dependence of the IAPs (positively or negatively). This could lead to strong spatial autocorrelation of the number of IAPs between cells and violate the assumptions of neutral models, suggesting the possible hotspots of high invasibility of certain areas in the KNP.

5.7 Optimal Sampling and Monitoring Effort

To determine the efficient sampling scheme, we first let s denote the total number of records reported (a sum of presences and absences), i.e. the sampling effort. For each unit of sampling effort, only one cell can be visited and the chance of reporting a presence record is equal to the proportion of presences of the cell. Five sampling schemes are examined, including random, systematic, addictive, elusive and random-walk. In a random sampling scheme, the cell visited each time is randomly chosen. In a systematic sampling scheme, all cells will be visited an equal number of times. In an addictive sample scheme, the ranger tends to visit the cells having

more presence records, i.e. the probability of visiting the cell i depends on its recording history (Eq. 5.7):

$$v(i)/\sum_{j} v(j) \tag{5.7}$$

where $v(i)$ and $v(j)$ are one plus the current number of presence records of cell i and j, respectively. Plus one is to ensure cells currently with no presence records can still have a chance to be visited. In an elusive sampling scheme, the ranger will try to avoid visiting the cells with presence records, with the probability of visiting cell i being (Eq. 5.8),

$$(1/v(i))/\sum_{j} (1/v(j)) \tag{5.8}$$

In a random-walk sampling scheme, the ranger will randomly choose a cell adjacent to the cell visited at the last time, i.e. randomly choosing among the four neighbouring cells of the current visiting cell. If the cell is at the border of KNP, the choice of adjacent cells will only be among those within the KNP.

We then simulated the sampling process according to the above schemes on a 4×4 km resolution for the whole KNP. Twenty-six simulations with different scenarios of sampling efforts ($s = 512, 724, 1024, \ldots, 3 \times 10^6$) were chosen for each random, addictive and random-walk schemes; 22 simulations with different sampling efforts (1333, 2666, ..., 3×10^6) for systematic scheme; 17 simulations with different sampling efforts (512, 724, 1024, ..., 131072) for elusive scheme. The maximum number of sampling effort was constrained by the computational capacity. Each simulation was then repeated five times to reduce the effect of stochasticity, and, thus, a total number of 585 simulations were run, with the maximum sampling effort in a single simulation reaching three million. The results from these simulations allowed us to further compare the performance from each sampling scheme and calculate the minimum sampling effort in the estimation of total abundance and the spatial structure of IAPs in the KNP.

The Poisson occupancy-abundance model explained a significant amount of variance in the relationship between abundance estimation and sampling effort (F-ratio, $p < 0.01$; Fig. 5.4a). In terms of accuracy, we compared the limit of the logarithmic abundance estimations ($\ln N_{\delta}$) with the AOO abundance estimations (High: $\ln(1450685) = 14.188$; Middle: $\ln(552820) = 13.22$; Low: $\ln(313461) = 12.66$). At least 13.7 % of current sampling effort (i.e. 3×10^5 records) is needed for reaching the 0.95-level accuracy when using the random sampling scheme. For detecting the spatial structure of IAPs in KNP (the spatial distribution of local abundance), the minimum sampling effort with satisfactory similarity ($d < d_{0.05}$) was attained at a mere level of 6.7×10^4 records for random sampling (2.78 % of current sampling effort; Fig. 5.4b). Current sampling efforts (121 records per km^2) could be reduced to 15.3 records per km^2 for an acceptable abundance estimate, and to 3.4 records per km^2 for distribution detection. Detecting the overall

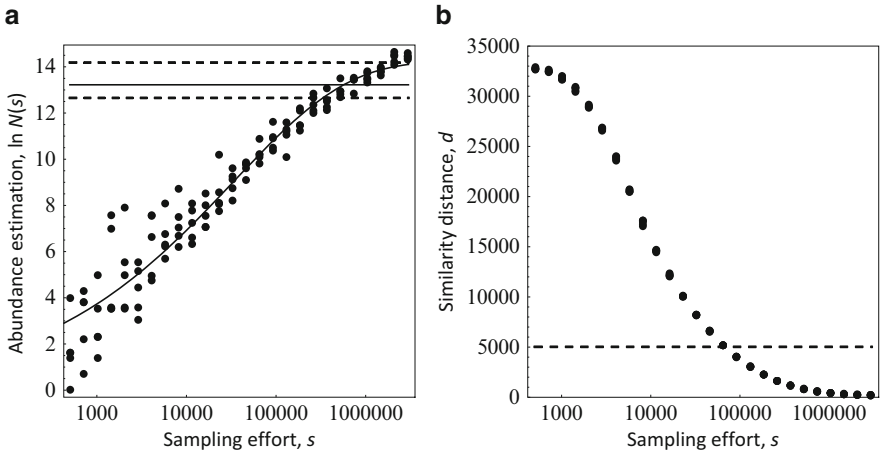

Fig. 5.4 (**a**) Abundance estimation and, (**b**) similarity distance as a function of sampling effort (s) for random sampling scheme. The *solid* and *dashed straight lines* in (**a**) correspond to the area-of-occupancy estimation of abundance (see Fig. 5.2a); the *dashed line* in (**b**) indicates the 0.05 critical level $d_{0.05}$ of the similarity distance. Similar plots for other sampling schemes are not shown due to resembling patterns

Table 5.1 Effects of sampling schemes and effort on detecting total abundance and spatial structure of invasive alien plants in Kruger National Park

Sampling scheme	$\ln N_\delta$	α	β	A_H	A_M	A_L	s^*	s^{**}
Random	14.24 ± 0.29	0.029	0.34	>0.99	0.93	0.89	15.77	3.46
Systematic	15.58 ± 0.45	0.059	0.26	0.91	0.85	0.81	20.94	3.44
Addictive	7.92 ± 0.10	0.002	0.81	0.21	0.33	0.40	–	–
Elusive	10.80 ± 0.36	0.007	0.55	0.69	0.78	0.83	–	3.45
Random-walk	15.16 ± 0.79	0.066	0.27	0.94	0.87	0.84	12.13	2.24

$\ln N_\delta$ indicates the converging limit of ln-transformed total abundance estimation; α and β are two model parameters; A_H, A_M and A_L the accuracy for abundance estimation with respect to the high, middle and low estimates; s^* indicates the minimum sampling effort (records per km^2) for $A_M = 0.95$; s^{**} is the minimum sampling effort (records per km^2) for the detected spatial structure not significantly different from observations ($d_s < d_{0.05}$); '–' indicates schemes failed to reach the 0.95 accuracy

spatial pattern of the IAP distribution takes much less effort than having an accurate assessment of abundance (Joseph et al. 2006). The accuracy A of the random sampling scheme showed consistently top ranking for abundance estimation, followed by random-walk, systematic, elusive and addictive (Table 5.1).

5.8 Cross-Scale Monitoring and Management

Efficient monitoring programmes for protected areas require robust methods for estimating target species abundance and distribution, and detecting changes thereof over time. Area-of-occupancy (AOO), using the scaling pattern of occupancy demonstrated here, fulfils the requirements of such a programme. Traditional mensuration methods of abundance estimating such as systematic or cluster sampling are only useful at local scales (e.g. 0.1–10 km^2 for complete counts) due to the method of data collection and costs. There is increasing interest in using binary (presence/absence) data for large scale surveys (Brotons et al. 2004; Joseph et al. 2006). In this regard, two categories of abundance estimation models have been developed. First, the intraspecific occupancy-abundance relationship is grounded in the ubiquitous positive correlation between species abundance and range size (Gaston and Blackburn 2000; He and Gaston 2003). Second, the scaling pattern of occupancy describes how adjacent occupied cells merge with increasing grain (Hartley and Kunin 2003; Hui et al. 2006; Lennon et al. 2007; Gaston and Fuller 2009). A multi-criteria test suggests the supremacy of the scaling pattern of occupancy models over the occupancy-abundance relationship models in estimating abundance and yielding macroecological patterns (Hui et al. 2009). Indeed, Kunin's (1998) power-law AOO requires only one-tenth of the current sampling effort for a robust estimate of IAP abundance in the KNP (Table 5.1). Furthermore, scaling pattern of occupancy models provide a framework for further analysing biodiversity patterns across scales (e.g. Hurlbert and Jetz 2007; Foxcroft et al. 2009). We thus suggest that both these methods are useful for estimating abundance and distribution in large-scale monitoring programmes in protected areas.

Even though the AOO method can capture the essence of species distributions across a range of scales, we found a slight change (decline) in the slope of the AOO, calibrated from large to small scales (Fig. 5.2). The box-counting fractal dimension also declines from 0.97 to 0.77, indicating a more scattered, less structured distribution at finer scales (Lennon et al. 2007). Two reasons for this are plausible. First, AOO only considers the scaling pattern of presence records, not absence records. The status of absence for cells at small scales is inferred by either without records or with a low number of absence records which result in an overestimation of absence cells in the small scaled grid cells (i.e. a low omission error; Pearce and Ferrier 2000; Anderson et al. 2003). This points to an underestimation of abundance when using small-scale AOO models due to the concave shape of the scaling pattern of occupancy. However, the high sampling intensity in the KNP can largely reduce the influence of such nonlinearity in the AOO that the power-law exponent and thus the fractal dimension declines when the spatial resolution increases (i.e. scaling down to finer scales). Second, this declining of fractal dimensions when scaling down reflects that the fractal structure of species distribution breaks down at small scales (Hartley et al. 2004; Hui and McGeoch 2007; Lennon et al. 2007). This, however, suggests an overestimation of abundance when using large-scale AOO. These two effects can compensate each other and allow for a reasonably robust estimate from

the AOO method. In addition, the fractal dimension at large scales could reflect the multidirectional range expansion through continuous habitat (Hui 2011). Furthermore, a fractal dimension close to one could indicate a linearised distribution of IAPs which can help to identify linear-shaped habitat (e.g. spread via rivers and roads) that determines the pathways of range expansion.

As with all methods for inference and extrapolation, the AOO method has its own limitations. First, uncertainty is often high for species with a low number of records. To have a fairly good estimate of abundance or coverage, we need at least 15–20 records per km^2. In addition, cells with a low number of records will also suffer from a high risk of a false categorization of non-detected species as absence. Other methods should thus also be consulted if this is suspected (e.g. MacKenzie et al. 2002; Hui et al. 2011a). Second, in general, species with a moderate range but which are locally abundant will have a high power-law exponent of the AOO across scales (i.e. steep scaling pattern of AOOs). However, the structures of species distribution across scales (e.g. at regional and local scales) are diverse (Gaston 1994). How these diverse structures of species distributions are related to the forms of AOOs across scales certainly needs more investigation. Finally, extrapolation across orders of magnitude carries inherent risks as each ecological process only works at a specific range of spatial scales. The power-law patterns can break up at very fine scales (e.g. Hui and McGeoch 2007). In this regard, extrapolation should only be used for scales close to the calibration range of the AOOs. However, this is often not possible for the purpose of rapid assessment of the invasion status. Managers have to base their decision on the trade-off between invasion risks and the uncertainty of inference. Extrapolation to the detection range (16×16 m) could be reliable (25 % of the AOO scale ranges from 125 m to 8 km in cell width) but the abundance estimates at 1 m^2 scale are crude and can only be used as an indicator for rapid risk assessment. Once the rapid assessment is done, managers can then choose more sophisticated approaches to investigate, and if needed, more detailed species-specific population-level structures and derive local-scale management plans.

In conclusion, to better manage biodiversity in protected areas, implementing knowledge gained in science into conservation management action is essential for conservation agencies. The use of the AOO model provides managers with a rigorous assessment of the options available for monitoring invasive alien plants over a large area.

Acknowledgments We are grateful to G. Blanchet, G. Cruz-Piñón, F. He, K. J. Gaston, I. Kühn, W. E. Kunin, S. Hartley and anonymous reviewers for comments and logistic help, and SANParks and the Centre for Invasion Biology, Stellenbosch University, for financial support. C.H., L.C.F and D.M.R. acknowledge financial support from the NRF. We thank the Kruger National Park rangers for collecting the CyberTracker data and Zuzana Sixtová for technical assistance with editing.

References

Anderson RP, Lew D, Peterson AT (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. Ecol Model 162:211–232

Brotons L, Thuiller W, Araújo MB et al (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography 27:437–448

Cohen JE, Jonsson T, Muller CB, Godfray HCJ, Savage VM (2005) Body sizes of hosts and parasitoids in individual feeding relationships. Proc Natl Acad Sci U S A 102:684–689

Dietemann V, Lubbe A, Crewe RM (2006) Human factors facilitating the spread of a parasitic honey bee in South Africa. J Econ Entomol 99:7–13

Du Toit JT, Rogers KH, Biggs HC (eds) (2003) The Kruger experience. Ecology and management of savanna heterogeneity. Island Press, Washington, DC

Foxcroft LC, Henderson L, Nichols GR et al (2003) A revised list of alien plants for the Kruger National Park. Koedoe 4:21–44

Foxcroft LC, Richardson DM, Rouget M et al (2009) Patterns of alien plant distribution at multiple spatial scales in a large national park: implications for ecology, management and monitoring. Divers Distrib 15:367–378

Gaston KJ (1994) Rarity. Chapman and Hall, London

Gaston KJ, Blackburn TM (2000) Pattern and process in macroecology. Blackwell Science, Oxford

Gaston KJ, Fuller RA (2009) The sizes of species' geographic ranges. J Appl Ecol 46:1–9

Halley JM, Hartley S, Kallimanis AS et al (2004) Uses and abuses of fractal methodology in ecology. Ecol Lett 7:254–271

Hartley S, Kunin WE (2003) Scale dependency of rarity, extinction risk, and conservation priority. Conserv Biol 17:1559–1570

Hartley S, Kunin WE, Lennon JJ et al (2004) Coherence and discontinuity in the scaling of species' distribution patterns. Proc R Soc B-Biol Sci 271:81–88

He F, Gaston KJ (2003) Occupancy, spatial variance, and the abundance of species. Am Nat 162:366–375

Hui C (2011) Forecasting population trend from the scaling pattern of occupancy. Ecol Model 222:442–446

Hui C, McGeoch MA (2007) Modeling species distributions by breaking the assumption of self-similarity. Oikos 116:2097–2107

Hui C, McGeoch MA (2008) Does the self-similarity species distribution model lead to unrealistic predictions? Ecology 89:2946–2952

Hui C, McGeoch MA, Warren M (2006) A spatially explicit approach to estimating species occupancy and spatial correlation. J Anim Ecol 75:140–147

Hui C, McGeoch MA, Reyers B et al (2009) Extrapolating population size from the occupancy-abundance relationship and the scaling pattern of occupancy. Ecol Appl 19:2038–2048

Hui C, Foxcroft LC, Richardson DM et al (2011a) Defining optimal sampling effort for large-scale monitoring of invasive alien plants: a Bayesian method for estimating abundance and distribution. J Appl Ecol 48:768–776

Hui C, Richardson DM, Robertson MP et al (2011b) Macroecology meets invasion ecology: linking the native distributions of Australian acacias to invasiveness. Divers Distrib 17:872–883

Hurlbert AH, Jetz W (2007) Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. Proc Natl Acad Sci U S A 104:13384–13389

Joseph LN, Field SA, Wilcox C et al (2006) Presence-absence versus abundance data for monitoring threatened species. Conserv Biol 20:1679–1687

Kruger JM, MacFadyen S (2011) Science support within the South African National Parks adaptive management framework. Koedoe 53(2):Art. #1010, 7 pages. doi:10.4102/koedoe.v53i2.1010

Kunin WE (1998) Extrapolating species abundance across spatial scales. Science 281:1513–1515

Lennon JJ, Kunin WE, Hartley S et al (2007) Species distribution patterns, diversity scaling and testing for fractals in southern African birds. In: Storch D, Marquet PA, Brown JH (eds) Scaling biodiversity. Cambridge University Press, Cambridge, pp 51–76

MacKenzie DI, Nichols JD, Lachman GB et al (2002) Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255

Parker IM, Simberloff D, Lonsdale WM et al (1999) Impact: toward a framework for understanding the ecological effects of invaders. Biol Invasive 1:3–19

Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. Ecol Model 133:225–245

Pickett STA, Cadenasso ML, Benning TL (2003) Biotic and abiotic variability as key determinants of savanna heterogeneity at multiple spatiotemporal scales. In: Du Toit JT, Rogers KH, Biggs HC (eds) The Kruger experience: ecology and management of savanna heterogeneity. Island Press, Seattle, pp 22–40

Preston FW (1948) The commonness, and rarity, of species. Ecology 29:254–283

Quinn GP, Keough MJ (2002) Experimental design and data analysis for biologists. Cambridge University Press, Cambridge

Royle JA, Nichols JD, Kery M (2005) Modelling occurrence and abundance of species when detection is imperfect. Oikos 110:353–359

Sokal RR, Rohlf FJ (1995) Biometry: the principles and practice of statistics in biological research. W. H. Freeman and Company, New York

Spear D, McGeoch MA, Foxcroft LC et al (2011) Alien species in South Africa's national parks. Koedoe 53:Art. #1032, 4 pages. doi:10.4102/koedoe.v53i1.1032

Volkov I, Banavar JR, He F et al (2005) Density dependence explains tree species abundance and diversity in tropical forests. Nature 438:658–661

Wang Z, Bovik AC, Sheikh HR et al (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13:600–612

Wilson RJ, Thomas CD, Fox R et al (2004) Spatial patterns in species distributions reveal biodiversity change. Nature 432:393–396

Wright DH (1991) Correlations between incidence and abundance are expected by chance. J Biogeogr 1:463–466